Developmental and Comparative Immunology 46 (2014) 300–313

Contents lists available at ScienceDirect

Developmental and Comparative Immunology

journal homepage: www.elsevier.com/locate/dci

Characteristics of the somatic hypermutation in the *Camelus dromedarius* T cell receptor gamma (TRG) and delta (TRD) variable domains

Salvatrice Ciccarese [a,*], Giovanna Vaccarelli [a], Marie-Paule Lefranc [b], Gianluca Tasco [c], Arianna Consiglio [d], Rita Casadio [c], Giovanna Linguiti [a], Rachele Antonacci [a]

[a] Department of Biology, University of Bari, via E. Orabona 4, 70125 Bari, Italy
[b] IMGT, Laboratoire d'ImmunoGénétique Moléculaire, Institut de Génétique Humaine, UPR CNRS 1142, Université Montpellier 2, 34396 Montpellier Cedex 5, France
[c] Biocomputing Group, CIRI-Health Science and Technologies/Department of Biology, University of Bologna, via Selmi 3, 40126 Bologna, Italy
[d] CNR, Institute for Biomedical Technologies of Bari, via Amendola, 70125 Bari, Italy

ARTICLE INFO

Article history:
Received 8 April 2014
Revised 5 May 2014
Accepted 5 May 2014
Available online 13 May 2014

Keywords:
T cell receptor
TRG and TRD genes
Somatic hypermutation
Dromedary
IMGT

ABSTRACT

In previous reports, we had shown in *Camelus dromedarius* that diversity in T cell receptor gamma (TRG) and delta (TRD) variable domains can be generated by somatic hypermutation (SHM). In the present paper, we further the previous finding by analyzing 85 unique spleen cDNA sequences encoding a total of 331 mutations from a single animal, and comparing the properties of the mutation profiles of dromedary TRG and TRD variable domains. The transition preference and the significant mutation frequency in the AID motifs (dgyw/wrch and wa/tw) demonstrate a strong dependence of the enzymes mediating SHM in TRG and TRD genes of dromedary similar to that of immunoglobulin genes in mammals. Overall, results reveal no asymmetry in the motifs targeting, i.e. mutations are equally distributed among g:c and a:t base pairs and replacement mutations are favored at the AID motifs, whereas neutral mutations appear to be more prone to accumulate in bases outside of the motifs. A detailed analysis of clonal lineages in TRG and TRD cDNA sequences also suggests that clonal expansion of mutated productive rearrangements may be crucial in shaping the somatic diversification in the dromedary. This is confirmed by the fact that our structural models, computed by adopting a comparative procedure, are consistent with the possibility that, irrespective of where (in the CDR-IMGT or in FR-IMGT) the diversity was generated by mutations, both clonal expansion and selection seem to be strictly related to an enhanced structural stability of the γδ subunits.

1. Introduction

Camelus dromedarius belongs to the suborder Tylopoda, is an atypical ruminant having only three of the four stomach cavities of ruminants and occupies in the Cetartiodactyls phylogeny a basal position compared with the other families of the "Ruminantia" suborder (Bininda-Emonds et al., 2007). The dromedary is known as an important immunological model for medical research since, in addition to the conventional tetrameric IgGs, it has special smaller heavy chain-only antibodies (Nguyen et al., 2001) which have been found to be particularly useful in biotechnological applications i.e. nanobiotechnologies (Muyldermans et al., 2009; Vincke et al., 2009; Deschacht et al., 2010). We have recently shown, for the first time in a mammalian organism, that somatic hypermutation (SHM) occurs in dromedary T cell receptor gamma (TRG) and delta (TRD) productively rearranged genes (Antonacci et al., 2011; Vaccarelli et al., 2012). Furthermore, SHM in TR chains has been postulated to occur in salmonids (Yazawa et al., 2008) and sharks (Chen et al., 2009, 2012). Therefore, our finding further differentiates from other ruminants where no indications have been found that SHM is able to generate repertoire diversity of TRG (Hein and Dudler, 1997; Vaccarelli et al., 2005; Herzig et al., 2006a) and TRD (Massari et al., 2000; Herzig et al., 2006b) V regions.

It is well-known that among mammals, SHM occurs primarily in germinal center B cells and is the driving force for antibody affinity maturation, since it introduces point mutations into the variable domains of immunoglobulin (IG) genes (Li et al., 2004). The dgyw/wrch motif (where d = a or g or t, y = c or t, w = t or a, r = a or g, and h = c or t or a) has been found to be the principal hotspot for AID-induced g:u lesions in rearranged IG genes during somatic hypermutation (Rogozin and Diaz, 2004; Liu and Schatz, 2009). In

* Corresponding author. Address: Department of Biology, University of Bari "Aldo Moro", via E. Orabona 4, 70125 Bari, Italy. Tel.: +39 080 5443338/84; fax: +39 080 5443386.
E-mail address: salvatricemaria.ciccarese@uniba.it (S. Ciccarese).

http://dx.doi.org/10.1016/j.dci.2014.05.001
0145-305X/

addition, the dinucleotide target wa/tw has been identified as a principal site for a/t mutations, which in turn has led to studies to define the roles of the error-prone polymerases in mismatch repair as secondary mutators (Pavlov et al., 2002; Zhao et al., 2013). In this paper we present a comprehensive study of SHM in dromedary γ/δ T cells. By analyzing 85 unique spleen cDNA transcripts from a single animal, we compare the properties of the mutation profiles of TRG and TRD variable domains. The main difference regards the mutation distribution which proves to be homogeneous throughout the entire length in framework regions (FR-IMGT) of TRD domains, whereas it is localized in both the complementarity determining regions (CDR-IMGT) and in the FR-IMGT in TRG domains. Overall, mutations observed in dromedary TRG and TRD are equally distributed among g:c and a:t base pairs and no strand bias is highlighted leading us to conclude that they could be largely dependent on the combined pathways of uracil-DNA glycosylase (UNG) and mismatch (MSH) repair.

Furthermore, analysis of clonal lineages suggests the occurrence of sustained and sequential mutational events and selection in both TRG and TRD cDNA clones. A more detailed analysis of clonal radiation in cDNA sequences showed that replacement (R) mutations in FRs appear to be early in the lineages. In addition, these R founder mutations are nonconservative of the amino acid physicochemical properties, i.e. most of them change non-hydrophobic amino acid residues to hydrophobic ones, implying that the selection for during their clonal expansion is favored by an increase in the structural intrinsic stability of the γδ receptor.

2. Materials and methods

2.1. The animal source of tissues

The *C. dromedarius* spleen and liver from a single animal were provided by Prof. Mohamed S. Hassanane (Cell Biology Department National Research Center, Dokki, Giza, Egypt).

2.2. Sequence analysis

We used 73 cDNA sequences from the spleen (Accession Nos. JF755949–JF755967; JF792633–JF792673; FN252369–FN252381) that were aligned with three known germlines: TRGV1 and TRGJ1 (Accession No. JN165102), TRGV2 and TRGJ2 (Accession No. JN172913), and TRDV4 including the TRDJ4 (Accession No. FN298231) genomic sequences from the liver, respectively. 12 cDNA TRDV1 clones (Accession Nos. FN252337–FN252338, FN252340–FN252344, FN252346–FN252350) from the above spleen, sharing the same VDJ rearrangement, were members of the same clonal set and were aligned with the obtained defined consensus (clone SC32) sequence (Accession No. FN252339) (Antonacci et al., 2011). The cDNA mutation analysis was conducted by *python script* (available upon request): the program detects, in a multiple alignment, and according to the reference sequence, the number of single-based and tandem substitutions per codon and classifies them as synonymous and nonsynonymous changes.

Nucleotide differences between aligned clones within each group (TRG and TRD) were principally single position differences located at various positions within the TRV encoded region. The number of clones within a clonal set ranged in size from 2 to 8 and identical as well as different nucleotide mutations were observed. By analyzing manually each single mutation in a total of 85 clones, 331 mismatches were identified in 39,819 TRV encoded nucleotides. Because a mutation could be carried during clonal expansion, a mutation observed in the same position in a clonal set was deemed to represent a single event and, therefore, was only counted once in the mutational analyses. The mutational analysis has been further confirmed following the model in sections 9 and 10 of the IMGT/V-QUEST Documentation (IMGT®, the international ImMunoGeneTics information system®, http://www.imgt.org) (Lefranc et al., 2009).

2.3. Identification of AID hot spots

A search for the number and the position of AID hot spots (a/g/t)g(c/t)(a/t) [or dgyw] and (a/t)(a/g)c(c/t/a) [or wrch] and wa/tw in germline TRGV and TRDV gene sequences was performed using the Scope Version 2.1.0 tool (http://genie.dartmouth.edu/scope/) (Carlson et al., 2007). Germline and cDNA sequences were multi-aligned using MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/) (Edgar, 2004) and the mutation profile for each nucleotide has been obtained for the TRGV1, TRGV2, TRDV4, and TRDV1 sequences, respectively. At the same time besides the dgyw/wrch motifs (Rogozin and Diaz, 2004), the search analysis for the number and the position of AID hot spots has been conducted on the standard motifs (a/g)g(c/t)(a/t) [or rgyw] and (a/t)(a/g)c(c/t) [or wrcy], following the model in section 11 of the IMGT/V-QUEST Documentation (IMGT®, the international ImMunoGeneTics information system®, http://www.imgt.org) (Lefranc et al., 2009).

2.4. Statistical analysis

Statistical analyses were performed using 2×2 contingency tables. All the *p*-values shown in the Results were obtained using the Chi-squared test, considering as statistically significant a *p*-value <0.05. Fisher's Exact test was used to confirm the significance of the Chi-squared test when the counts of observed samples had values <5. When performing multiple comparisons among mutation counts in different regions, the Chi-squared test *p*-values were adjusted using Benjamini–Hochberg false discovery rate (Benjamini and Hochberg, 1995). All the analyses were performed using the R software environment for statistical computing (http://www.r-project.org/).

2.5. Assessment of antigen selection from mutation profiles

The presence of antigen selection and affinity maturation was determined by comparing the R/S ratios between the CDR-IMGT and FR-IMGT using chi-squared analysis. All the *p*-values were computed with the Chi-squared test and confirmed with Fisher's Exact test.

2.6. Protein modeling

Modeling of domains obtained by joining AA sequences of germline V and J, TRGV1-TRGJ1-1 (TRGV1), TRGV2-TRGJ2-2 (TRGV2), TRDV4-TRDJ4 (TRDV4), and of domains of mutated cDNA clones, RTS88 (Accession No. JF755952), 5R1S169 (Accession No. JF792640), and RTVD4m14 (Accession No. FN252376) was carried out adopting the building by the homology procedure. The template was selected from the Protein Data Bank (PDB) on the basis of sequence/function similarity with the target sequence and it was the human γδ T cell receptor solved with an atomic resolution of 3Å (PDB code: 3omz) (Ehrenmann et al., 2010; Xu et al., 2011). The template sequences share 43%, 34%, and 35% sequence identity with TRGV1, TRGV2, and TRDV4, respectively. Global alignment of the target and template sequences was performed with ClustalW (http://www.ebi.ac.uk/clustalw/index.html) (Larkin et al., 2007). Furthermore, when necessary, alignment was manually adjusted after predicting the secondary structure of the target and aligned to that of the template as derived with the DSSP program (Kabsch and Sander, 1983). The secondary structure prediction